

Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2002, 11:51:58 ; Search time 50.08 Seconds

(without alignments)  
40.891 Million cell updates/sec

Title: US-09-432-546-5  
Perfect score: 103  
Sequence: 1 SRPMPWMPWKPLI 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SPREMBL\_17:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mhc:\*  
9: sp\_organelle:\*  
10: sp\_phage:\*  
11: sp\_plant:\*  
12: sp\_rodent:\*  
13: sp\_virus:\*  
14: sp\_vertebrate:\*  
15: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

## BEST AVAILABLE COPY

Result No.	Score	Query Match	Length	DB ID	Description
1	57	55.3	1662	2 P71431	P71431 Leptothrix
2	56	54.4	236	2 O9KFN3	O9KFN3 bacillus ha
3	55.5	53.9	165	10 O9SNN3	O9SNN3 oryza sativ
4	55.5	53.9	676	10 O9FG26	O9FG26 arbidopsi
5	53.5	51.9	491	4 O9BU12	O9BU12 homo sapien
6	53	51.5	95	10 O9LQNO	O9LQNO arbidopsi
7	52.5	51.0	970	10 O80790	O80790 arbidopsi
8	52.5	51.0	1231	10 O82276	O82276 arbidopsi
9	52.5	50.5	157	5 O9YOE8	O9YOE8 drosophila
10	52	50.5	301	2 O915M4	O915M4 pseudomona
11	52	50.5	1245	3 O9Y7V5	O9Y7V5 trichoderma
12	51.5	50.0	109	10 O40548	O40548 nicotiana t
13	51.5	50.0	196	10 O08195	O08195 nicotiana t
14	51.5	50.0	209	10 O08194	O08194 nicotiana t
15	51.5	50.0	321	10 O9SJZ8	O9SJZ8 arbidopsi
16	51.5	50.0	2970	12 O56073	O56073 hepatitis g
17	51	49.5	1173	12 O990M4	O990M4 human coron
18	51	49.5	1173	12 O990M3	O990M3 human coron
19	51	49.5	1173	12 O990M2	O990M2 human coron

20	51	49.5	1173	12 O990M1	O990M1 human coron
21	51	49.5	1383	12 O84712	O84712 porcine epi
22	50.5	49.0	141	11 O9CCAL	O9CCAL mus musculu
23	50.5	49.0	560	5 O44626	O44626 caenorhabdi
24	50	48.5	83	2 O9WYF1	O9WYF1 thermotoga
25	50	48.5	492	10 O9AV15	O9AV15 oryza sativ
26	49.5	48.1	257	2 O56924	O56924 yersinia en
27	49.5	48.1	723	12 O9DHC4	O9DHC4 tt virus. o
28	49	47.6	159	2 O9KZT3	O9KZT3 streptomyce
29	49	47.6	169	1 O9Y9Q4	O9Y9Q4 aeropyrum p
30	49	47.6	251	2 O52853	O52853 bacillus su
31	49	47.6	467	5 O19573	O19573 caenorhabdi
32	49	47.6	485	2 O9AFK1	O9AFK1 shigella fl
33	49	47.6	669	5 O18300	O18300 caenorhabdi
34	49	47.6	746	12 O9JH31	O9JH31 tt virus. o
35	49	47.6	1053	4 O9BXK0	O9BXK0 homo sapien
36	49	47.6	1487	5 O9VLP5	O9VLP5 drosophila
37	48	46.6	196	10 O9LP00	O9LP00 arbidopsi
38	48	46.6	236	4 O9POT8	O9POT8 homo sapien
39	48	46.6	314	10 O9ZUY7	O9ZUY7 arbidopsi
40	48	46.6	947	10 O9IDL8	O9IDL8 arbidopsi
41	48	46.6	1172	10 O9LPO5	O9LPO5 arbidopsi
42	47.5	46.1	111	5 O18753	O18753 caenorhabdi
43	47.5	46.1	752	2 O9KOW8	O9KOW8 vibrio chol
44	47	45.6	73	2 O9RM06	O9RM06 deinococcus
45	47	45.6	133	5 O9VX67	O9VX67 drosophila

## ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	1662 AA.
P71431				
AC	P71431			
DT	01-FEB-1997 (TREMBL)			
DT	01-FEB-1997 (TREMBL)			
DT	01-JUN-2001 (TREMBL)			
DE	MORA GENE ENCODING MANGANESE OXIDIZING PROTEIN PRECURSOR.			
GN	MORA.			
OC	Leptothrix discophora.			
OC	Bacteria; Proteobacteria; beta subdivision; Comamonadaceae;			
OC	Leptothrix.			
OX	NCBI_TaxID=89;			
RN	[1]			
RP	SEQUENCE OF 1-1150 FROM N.A.			
RC	STRAIN=SS-1;			
RC	STRAIN=SS-1;			
RA	Corsjens P.L.;			
RL	Thesis (1993), Biochemistry, Leiden University, The Netherlands.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SS-1;			
RA	Corsjens P.L.;			
RL	Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SS-1;			
RA	Corsjens P.L.A.M., de Vrind J.P.M., Goosen T., de Vrind-de Jong E.W.;			
RL	Geomicrobiol. J. 14:91-106(1997).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SS-1;			
RA	Corsjens P.L.;			
RL	Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; 225774; CAAB1037.1; -			
DR	InterPro; IPR001865; Ribosomal_S2.			
DR	PROSITE; PS00962; RIBOSOMAL_S2_1; UNKNOWN_1.			
FT	SIGNAL			
FT	CHARIN			
FT	SEQUENCE			

Query Match 55.3%; Score 57; DB 2; Length 1662;  
 Best Local Similarity 63.6%; Pred. No. 16;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 3 RMPWPKWPL 13  
 |||||  
 DB 490 RMPWPKWPL 500

## RESULT 2

O9KFN3 PRELIMINARY; PRT; 236 AA.  
 AC O9KFN3;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE BH0446 PROTEIN.  
 GN BH0446.  
 OS Bacillus halodurans.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 CX NCBI\_TaxID=86665;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C-125 / JCM 9153;  
 RX MEDLINE=20512582; PubMed=11058132;  
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
 RA Fuji F., Hirata C., Nakamura Y., Ogasawara N., Kuhara S.,  
 RA Horikoshi K.;  
 RT \*Complete genome sequence of the alkaliphilic bacterium Bacillus  
 RT halodurans and genomic sequence comparison with Bacillus subtilis.\*;  
 RL Nucleic Acids Res. 28:4317-4331(2000).  
 DR EMBL: AP001508; BAB04165.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 236 AA; 2698 MW; A0BD6DD9EFE34FA9 CRC64;

Query Match 54.4%; Score 56; DB 2; Length 236;  
 Best Local Similarity 77.8%; Pred. No. 3.7;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 6 WMPWKWPL 14  
 |||||  
 DB 192 WMPWKWPL 200

## RESULT 3

O9SNN3 PRELIMINARY; PRT; 165 AA.  
 AC O9SNN3;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
 DE HYPOTHETICAL PROTEIN.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Euryaliaceae; Oryzaceae; Oryza.  
 CX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nippondare(GAS) genomic DNA, chromosome 6, PAC  
 RT clone:PO493C11."  
 RT Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP000559; BAA84796.1; -.  
 SQ SEQUENCE 165 AA; 282558BD22DE9 CRC64;

DB 10; Length 165;  
 1;  
 0; Indels 1; Gaps 1;

OY 5 PMPWPKW 12  
 |||||  
 DB 40 PMPWPKW 48

## RESULT 4

O9FG26 PRELIMINARY; PRT; 676 AA.  
 AC O9FG26;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE NON-LTR RETROELEMENT REVERSE TRANSCRIPTION-LIKE.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
 CX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COLDBIA;  
 RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,  
 RA Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI."  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP002032; BAB09815.1; -.  
 KW RNA-directed DNA polymerase.  
 SQ SEQUENCE 676 AA; 7665 MW; AD1C1E18E46F53A CRC64;

Query Match 53.9%; Score 55.5; DB 10; Length 676;  
 Best Local Similarity 47.1%; Pred. No. 11;  
 Matches 8; Conservative 0; Mismatches 2; Indels 7; Gaps 1;

OY 2 RRPWPKW 11  
 |||||  
 DB 443 RSWPTLFTLVWGMKW 459

## RESULT 5

O9BU12 PRELIMINARY; PRT; 491 AA.  
 AC O9BU12;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE SIMILAR TO FERREDOXIN REDUCTASE.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 CX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LUNG CARCINOMA;  
 RA Strausberg R.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC002960; AAH02960.1; -.  
 SQ SEQUENCE 491 AA; 53836 MW; 15D07E714F592C9D CRC64;

Query Match 51.9%; Score 53.5; DB 4; Length 491;  
 Best Local Similarity 61.5%; Pred. No. 15;  
 Matches 8; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

OY 1 SRRPMPWK-WP 12  
 |||||  
 DB 3 SRCWRWGMWSAMP 15

RESULT 6  
 O9LONO PRELIMINARY; PRT; 95 AA.

RP SEQUENCE FROM N.A.  
RC STRAIN-CV, COLUMBIA;  
RA Theologis A.;  
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV, COLUMBIA;  
RA Theologis A.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV, COLUMBIA;  
RA Theologis;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
RN EMBL: AC007767; AAF81325.1; -  
SQ SEQUENCE 95 AA; 10624 MM; 0C33985771EBB54E CRC64;

RESULT	7
080790	
ID	080790
AC	080790;
DT	01-NOV-1998 (TREMBLrel. 08, Created)
DT	01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE	REVERSE-TRANSCRIPTASE-LIKE PROTEIN.
NCBI	

```

CC -| SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE
CC TRANSCRIPTASE).
DR EMBL: AC004483; AAC26674.1; -.
DR InterPro: IPR000477; RVTse.
DR Pfam: PF00078; rvt; 2.
KW RNA-directed DNA polymerase.
SQ SEQUENCE 970 AA; 110578 MW; 3CA514F63BF67FA7 CRC64;

Query Match 51.0%; Score 52.5; DB 10; Length 970;
Best Local Similarity 46.7%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 1; Indels 7; Gaps 1

QY 4 WP-----WMPWKW 11
|||
DB 739 WPTLFGMGIMWAKW 753

RESULT 8
082276 PRELIMINARY; PRT; 1231 AA.
ID 082276
AC 082276;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PUTATIVE REVERSE TRANSCRIPTASE.
GN T16B12.11.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eustroids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
RA Shen M., Romling C.M., Fraser C.M., Somerville C.R., Venter J.C.;
RT "Arabidopsis thaliana chromosome II BAC T16B12 genomic sequence.";
RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.
CC -| SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE
CC TRANSCRIPTASE).
CC EMBL: AC005311; AAC63844.1; -.
DR InterPro: IPR000477; RVTse.
DR InterPro: IPR000169; Thiolprot_act_site.
DR Pfam: PF00078; rvt; 1.
DR PROSITE: PS00639; THIOL_PROTEASE_HIS; UNKNOWN_1.
KW RNA-directed DNA polymerase.
SQ SEQUENCE 1231 AA; 139476 MW; 5832FD600F342BB1 CRC64;

Query Match 51.0%; Score 52.5; DB 10; Length 1231;
Best Local Similarity 46.7%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 1; Indels 7; Gaps 1;

QY 4 WP-----WMPWKW 11
|||
DB 1000 WPTLFGMGIMWAKW 1014

RESULT 9
09Y0E8 PRELIMINARY; PRT; 157 AA.
ID 09Y0E8
AC 09Y0E8;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SALIVARY GLAND SECRETION PROTEIN (FRAGMENT).
DE SGSI OR SGS-1 OR CG3047.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;

```

RN SEQUENCE FROM N.A.  
 RA Roth G.E., Matlier S., Bornschein H., Lehmann M., Korge G.;  
 RT "Structure and regulation of the salivary gland secretion protein gene  
 RT Sgs-1 of *Drosophila melanogaster*.";  
 RL Genetics 0:0-0(1999).  
 DR EMBL: AF156228; AAD43808.1; -  
 DR FLYBASE: FBgn0003372; Sgs1.  
 DR InterPro: IPR002965; P-rich\_extensn.  
 DR PRINTS: PR01217; PRICHEXTNSN.  
 FT NON\_TER  
 SQ SEQUENCE 157 AA; 17975 MW; 8495D86262819707 CRC64;

Query Match 50.5%; Score 52; DB 5; Length 157;  
 Best Local Similarity 75.0%; Pred. No. 8.1;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 PWWPFWKWP 12  
 Db 114 PWWPFWKWP 121

RESULT 10  
 Q915M4 PRELIMINARY; PRT; 301 AA.  
 AC Q915M4;  
 DT 01-MAR-2001 (TREMblrel. 16, Created)  
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE HISTOETHICAL PROTEIN PA0702.  
 GN PA0702.  
 OS Pseudomonas aeruginosa.  
 CC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 CC Pseudomonas.  
 CC NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PA01;  
 RC MEDLINE=20437337; PubMed=10964043;  
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Gader R.L., Goltner L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig R., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an  
 RT opportunistic pathogen.";  
 RL Nature 406:959-964(2000).  
 DR EMBL: AE004506; AAC04091.1; -  
 DR InterPro: IPR001541; Sterol\_desat.  
 DR Hypothetical protein; Complete proteome.  
 KW SEQUENCE 301 AA; 33816 MW; 09FC2935F490ABD7 CRC64;

Query Match 50.5%; Score 52; DB 2; Length 301;  
 Best Local Similarity 70.0%; Pred. No. 15;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 WPMWPMKWP 13  
 Db 97 WPMWPMKWP 106

RESULT 11  
 Q9Y7V5 PRELIMINARY; PRT; 1245 AA.  
 AC Q9Y7V5;  
 DT 01-NOV-1999 (TREMblrel. 12, Created)  
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)  
 GN CANDIDOSPORE SURFACE PROTEIN.  
 DE Cmpl.

OS Trichoderma harzianum.  
 CC Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Trichoderma.  
 CC NCBI\_TaxID=5544;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 32173;  
 RA Puyesky M., Benhamou N., Ponce Noyola P., Bauw G., Ziv T.,  
 RA van Montagu M., Herrera Estrella A., Horvitz B.A.;  
 RT "Developmental regulation of a gene encoding a multidomain  
 RT candidospore surface protein of *Trichoderma* cml.";  
 RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AJ133651; CAB40845.1; -  
 SQ SEQUENCE 1245 AA; 135824 MW; 3249C749AFA0C0F8 CRC64;

Query Match 50.5%; Score 52; DB 3; Length 1245;  
 Best Local Similarity 62.5%; Pred. No. 52;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 WPMWPMKWP 11  
 Db 1199 WPMWPMKWP 1206

RESULT 12  
 Q40548 PRELIMINARY; PRT; 109 AA.  
 ID Q40548;  
 AC Q40548;  
 DT 01-JAN-1998 (TREMblrel. 05, Created)  
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE PISTIL-SPECIFIC EXTENSIN-LIKE PROTEIN PRECURSOR (FRAGMENT).  
 OS Nicotiana tabacum (Common tobacco).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 CC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
 CC NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. PETITE HAVANA; TISSUE=PISTIL;  
 RC MEDLINE=93005740; PubMed=1392607;  
 RA Goldman S., Pezzotti M., Saurin J., Mariani C.;  
 RT "Developmental expression of tobacco pistil-specific genes encoding  
 RT novel extensin-like proteins.";  
 RL Plant Cell 4:1041-1051(1992).  
 CC -1- TISSUE SPECIFICITY: PISTIL (STIGMA AND STYLE TISSUE).  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION BEGINS IN FLORAL BUDS AFTER PISTIL  
 CC DIFFERENTIATION AND LEVELS GRADUALLY INCREASE DURING FLOWER  
 CC DEVELOPMENT TOWARD ANTHERESIS. LEVELS GRADUALLY DECREASE AFTER  
 CC POLLINATION AND ARE ABSENT BY THE SIXTH DAY AFTER POLLINATION.  
 DR EMBL: Z14014; CAA78392.1; -  
 DR Mendel; 16906; Nicotiana; 2747; 16906.  
 DR InterPro: IPR002965; P-rich\_extensn.  
 DR PRINTS: PR01217; PRICHEXTNSN.  
 KW structural protein; Repeat; Glycoprotein; Signal.  
 FT SIGNAL 1 22  
 FT CHAIN 23 109  
 FT DOMAIN 81 109  
 FT REPEAT 81 85  
 FT REPEAT 93 97  
 FT REPEAT 105 109  
 FT NON\_TER 109 109  
 SQ SEQUENCE 109 AA; 12373 MW; 6A6BCFAC192AB8C CRC64;

Query Match 50.0%; Score 51.5; DB 10; Length 109;  
 Best Local Similarity 43.8%; Pred. No. 6.8;  
 Matches 7; Conservative 1; Mismatches 1; Indels 7; Gaps 1;

QY 4 WPMWPMKWP 12  
 Db 43 WPMWPMKWP 58

RESULT 13  
ID 008195 PRELIMINARY; PRT; 196 AA.  
AC 008195;  
DT 01-NOV-1996 (TREMblrel. 01, Created)  
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)  
DE CYSTEINE-RICH EXTENSIN-LIKE PROTEIN 2.  
OS Nicotiana tabacum (Common tobacco).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
NCBI\_TaxID=4097;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93342083; PubMed=8341705;  
RA Wu H.M., Zou J., May B., Gu O., Cheung A.Y.;  
RT "A tobacco gene family for flower cell wall proteins with a proline-  
rich domain and a cysteine-rich domain."  
Proc. Natl. Acad. Sci. U.S.A. 90:6829-6833(1993).  
EMBL: L13440; AAA34060.1; -  
DR Mendel: 16902; Nicta:2747;16902.  
SQ SEQUENCE 196 AA; 21913 MW; 4C44E23C5B706E30 CRC64;

Query Match 50.0%; Score 51.5; DB 10; Length 196;  
Best Local Similarity 43.8%; Pred. No. 11;  
Matches 7; Conservative 1; Mismatches 1; Indels 7; Gaps 1;

Y 4 WPW-----WPWKW 12  
||| ||| ||  
Db 40 WPWEIPCYLTWPPWP 55

RESULT 14  
ID 008194 PRELIMINARY; PRT; 209 AA.  
AC 008194;  
DT 01-JAN-1998 (TREMblrel. 05, Created)  
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)  
DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)  
DE PISTIL-SPECIFIC EXTENSIN-LIKE PROTEIN PRECURSOR  
(CYSTEINE-RICH EXTENSIN-LIKE PROTEIN 1).  
OS Nicotiana tabacum (Common tobacco).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
NCBI\_TaxID=4097;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93342083; PubMed=8341705;  
RA Wu H.M., Zou J., May B., Gu O., Cheung A.Y.;  
RT "A tobacco gene family for flower cell wall proteins with a proline-  
rich domain and a cysteine-rich domain."  
Proc. Natl. Acad. Sci. U.S.A. 90:6829-6833(1993).  
EMBL: L13440; AAA34060.1; -  
DR Mendel: 16902; Nicta:2747;16902.  
SQ SEQUENCE 209 AA; 36192 MW; 43736712301D41BA CRC64;

Query Match 50.0%; Score 51.5; DB 10; Length 321;  
Best Local Similarity 46.7%; Pred. No. 18;  
Matches 7; Conservative 0; Mismatches 1; Indels 7; Gaps 1;

Y 4 WP-----WPWKW 11  
||| ||| |||  
Db 90 WPTVEVMAVWGMW 104

FT SIGNAL 1 19  
FT CHAIN 20 209  
FT DOMAIN 81 126  
FT REPEAT 81 85  
FT REPEAT 93 97  
FT REPEAT 105 109  
FT REPEAT 117 121  
FT REPEAT 122 126  
FT CARBOHYD 146 146  
SQ SEQUENCE 209 AA; 23351 MW; 12198BE2B8E08ED5 CRC64;

Query Match 50.0%; Score 51.5; DB 10; Length 209;  
Best Local Similarity 43.8%; Pred. No. 12;  
Matches 7; Conservative 1; Mismatches 1; Indels 7; Gaps 1;

Y 4 WPW-----WPWKW 12  
||| ||| |||  
Db 45 WPWEIPCYLTWPPWP 60

RESULT 15  
ID 09S328 PRELIMINARY; PRT; 321 AA.  
AC 09S328;  
DT 01-MAY-2000 (TREMblrel. 13, Created)  
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)  
DE PUTATIVE NON-LTR RETROELEMENT REVERSE TRANSCRIPTASE.  
GN ATG622350.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20083487; PubMed=10617197;  
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,  
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,  
RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,  
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,  
RA Copenhaver G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,  
RA Salzberg S.L., Fraser C.M., Venter J.C.;  
RT "Sequence and analysis of chromosome II of Arabidopsis thaliana."  
Nature 402:761-768(1999).  
EMBL: AC006592; AAD22368.1; -  
DR Nature: 402:761-768(1999).  
SQ SEQUENCE 321 AA; 36192 MW; 43736712301D41BA CRC64;

Query Match 50.0%; Score 51.5; DB 10; Length 321;  
Best Local Similarity 46.7%; Pred. No. 18;  
Matches 7; Conservative 0; Mismatches 1; Indels 7; Gaps 1;

Y 4 WP-----WPWKW 11  
||| ||| |||  
Db 90 WPTVEVMAVWGMW 104

Search completed: January 30, 2002, 11:51:58  
Job time: 197 sec

